UNIVERSIDAD AUTÓNOMA DE MADRID

ESCUELA POLITÉCNICA SUPERIOR

Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación

TRABAJO FIN DE GRADO

Estudio de Algoritmos Automáticos para Detección de Atributos Faciales y Soft Biometrics

Autor: Pablo Vicente Moñivar
Tutor: Ester González Sosa
Ponente: Julián Fiérrez Aguilar

JULIO 2017

Estudio de Algoritmos Automáticos para Detección de Atributos Faciales y Soft Biometrics

AUTOR: Pablo Vicente Moñivar
TUTOR: Ester González Sosa

Biometric Recognition Group - ATVS
Dpto. de Tecnología Electrónica y de las Comunicaciones
Escuela Politécnica Superior
Universidad Autónoma de Madrid
Julio 2017

Resumen

Desbloquear tu teléfono móvil, acceder a un recinto o hacer una transacción en el banco son algunas de las muchas acciones cotidianas que gracias a la biometría se han vuelto más cómodas y rápidas. Así pues, la biometría, puede definirse como la disciplina que se encarga del estudio de los rasgos físicos y de comportamiento con el objetivo del reconocimiento de personas.

La gran mayoría de los sistemas de reconocimiento se centra en el análisis de características propias de cada persona. La retina, el iris, la cara o las huellas dactilares son ejemplos de características físicas, mientras que la forma de andar o de escribir serían ejemplos de características dinámicas.

Sin embargo, hay un conjunto de características propias de cada persona, como el género, la edad, la altura o el peso, que proporcionan alguna información sobre la identidad, pero que de por sí solas son insuficientes para lograr determinar dicha identidad. Estas características se denominan Soft Biometrics. Y aunque los Soft Biometrics por si solos no permiten detectar una identidad, el conocimiento de ellos puede permitir una mejora del rendimiento en el reconocimiento de personas. Así, ventajas como una mayor discriminación en el rango de búsqueda o una mayor fiabilidad en las comparaciones obtenidas permiten la fabricación de sistemas de reconocimiento biométrico más avanzados.

En este Trabajo de Fin de Grado se ha realizado un estudio del rendimiento de dos sistemas comerciales destinados a la estimación automática de Soft Biometrics analizando su rendimiento y robustez a través de diferentes bases de datos.

Una vez realizado dicho estudio, se ha explorado la utilidad de los Soft Biometrics como complemento a los sistemas de reconocimiento facial. Para ello se han empleado los resultados obtenidos de los dos sistemas de estimación automática junto con el resultado proporcionado por un sistema de estimación manual.

Palabras clave

Soft Biometric, reconocimiento biométrico, detección, identificación, estimación

Abstract

Unlocking your mobile phone, accessing an enclosure or doing a transaction in the bank are some of the many everyday actions that thanks to biometrics have become more comfortable and faster. Thus, biometrics can be defined as the discipline that is responsible for the study of physical and behavioral traits with the aim of recognizing people.

The vast majority of recognition systems focuses on the analysis of each person's own features. The retina, the iris, the face or the fingerprints are examples of physical characteristics, while the way of walking or writing would be examples of dynamic features.

However, there are a set of characteristics of each person, such as gender, age, height or weight, which provide some information about identity, but alone are insufficient to determine that identity. These characteristics are called Soft Biometrics. And while Soft Biometrics alone cannot detect an identity, knowledge of them can allow improved performance in recognizing people. Thereby, advantages such as greater discrimination in the search range or greater reliability in the comparisons obtained, allow the manufacture of more advanced biometric recognition systems.

In this Bachelor Thesis, a study of the performance of two commercial systems for the automatic estimation of Soft Biometrics has been carried out, analyzing its performance and robustness through different databases.

Once the study was carried out, the use of Soft Biometrics was explored as a complement to facial recognition systems. For this purpose, we have used the results obtained from the two automatic estimation systems together with the result provided by a manual estimation system.

Keywords

Soft Biometric, biometric recognition, detection, identification, estimation

Agradecimientos

La realización y entrega de este Trabajo no se concibe como la mera presentación de un documento de investigación, sino como el final de un bonito ciclo de mi vida, un ciclo en que me he desarrollado a nivel personal, social y educacional, un ciclo en el que he estado acompañado de muchas personas que me han ayudado, apoyado y repartido cariño y que por ello quería agradecérselo.

En primer lugar, deseo expresar mi agradecimiento a toda mi familia, y en especial a mis padres y hermanos, con los que he vivido grandes momentos y alguno que otro triste como esas dos finales europeas que perdió nuestro Atleti. Su apoyo ha sido siempre incondicional, y sin ellos nunca habría podido llegar hasta este punto. Muchas gracias de corazón.

Por otro lado, quería agradecer especialmente a mi pareja, Mercedes, por todo el cariño que me ha dado y por haberme soportado, que no es fácil. Y a mis amigos de toda la vida Julio, Juan, David, Jaime, Pablo y Pedro, que nunca han dejado de confiar en mí. También quería mencionar a mis amigos de la universidad, especialmente a mis amigos Rodrigo, Beltrán, Carlos y Víctor, por hacerme disfrutar de unos estupendos años que nunca se olvidarán.

Por último, me gustaría agradecer la colaboración y el apoyo de la que ha sido mi tutora, Ester, que junto con el equipo de BiDA - ATVS me brindó la oportunidad de realizar este trabajo e introducirme en este mundo tan apasionante como el de la biometría facial en un año que sé que ha sido muy importante para ella.

Muchas gracias a todos.

INDICE DE CONTENIDOS

INDICE DE FIGURAS

INDICE DE TABLAS

1 Introducción

1.1 Motivación

Todos los días, para entrar en casa, desbloquear el móvil o hacer una transacción con el banco necesitas de una llave o clave. La facilidad de perder u olvidar éstas provoca que te quedes en una situación complicada, y que en caso de que otra persona encuentre la llave o descubra la contraseña, tenga la facilidad de utilizarlas como si fueras tú. Gracias a los sistemas biométricos todo esto está cambiando, las llaves y contraseñas se están sustituyendo por rasgos característicos de cada una de las personas como son su cara o huella dactilar. Estos sistemas no solo dan la comodidad de no tener que aprender o guardar ninguna cosa, sino que aumentan el nivel de seguridad.

El reconocedor facial es probablemente el más usado de todos los reconocedores biométricos de hoy en día, se puede ver en aeropuertos, estaciones o en los teléfonos móviles, y dado que su principal objetivo es la seguridad, su rendimiento y precisión tienen que ser muy altos. Tradicionalmente, el reconocimiento facial se desglosa principalmente en dos enfoques: atender la imagen de entrada a través de puntos de referencia faciales correspondientes a las diferentes regiones faciales como la nariz, boca u ojos o enfocar la imagen como un punto de referencia global de donde extraer las características para el reconocimiento. Pero el uso de otra información auxiliar en el reconocimiento facial es un campo todavía poco explorado. Los Soft Biometrics son aquellas características propias de cada individuo que se extraen del cuerpo humano, como por ejemplo la edad, el género, el color de los ojos o la altura, que son fácilmente distinguibles en la distancia y que perduran con el tiempo, pero que, de por sí solas son insuficientes para la identificación de un individuo. Sin embargo, la información que los Soft Biometrics proporciona puede resultar muy útil si se mezcla con un reconocedor biométrico, proporcionando mejores resultados tanto en el reconocimiento como en la discriminación del rango de búsqueda.

Durante este Trabajo de Fin de Grado (TFG), se pretende analizar dos potentes detectores de atributos faciales y con los resultados obtenidos y los rendimientos de cada uno analizados se pretende observar la influencia que puede tener un Soft Biometric en el reconocimiento de una persona.

De obtenerse unos buenos resultados estos podrían considerarse muy interesantes en todo el ámbito biométrico del reconocimiento facial y en otras aplicaciones como la seguridad, pudiendo ser utilizadas estas tecnologías a la hora de, por ejemplo, detectar a una persona en un aeropuerto sabiendo únicamente algún rasgo característico suyo como sería el tener barba y ser de etnia blanca.

1.2 Objetivos

El principal objetivo de este TFG, como se puede observar por el título, es llevar a cabo un estudio sobre algoritmos automáticos para detección de atributos faciales y Soft Biometrics. Para poder lograr ese objetivo ha sido necesario dividir y completar los siguientes objetivos específicos:

- El primer objetivo consistía en el manejo de grandes bases de datos de imágenes, como han sido las usadas durante este trabajo, y ver la calidad tanto de sus imágenes como de sus etiquetados para poder realizar una buena investigación. Para cumplir este objetivo en la primera parte del trabajo se llevó a cabo un análisis de cada una de las bases de datos utilizadas.
- Una vez obtenidas bases de datos fiables, el segundo objetivo consistía en aplicar a dichas bases los dos sistemas de estimación de atributos utilizados durante el trabajo con el fin de poder estudiar dichos sistemas.
- Otro de los objetivos más importantes fue el de realizar un sistema de que fusionaba la información obtenida por los estimadores de atributos con un sistema de reconocimiento facial y ver en los resultados, el efecto que provocaba la adicción de información relativa a los Soft Biometrics en el reconocimiento facial. Para este objetivo se necesitó haber realizado previamente los otros objetivos de análisis de la base de datos y de los sistemas de detección porque se usarían datos obtenidos de ellos.

Con la unión de todos estos objetivos se ha podido cumplir el objetivo principal y realizar un análisis conciso de diferentes sistemas de detección de atributos faciales viendo la importancia que tiene un Soft Biometric en un sistema de este calibre.

1.3 Metodología y planificación

El seguimiento que se llevó a cabo para la realización de este Trabajo de Fin de Grado comenzó en junio de 2016.

Inicialmente se llevó a cabo una toma de contacto y formación con todo el sistema de estimación de atributos, detección de caras y entendimiento del concepto Soft Biometric, para poder luego aplicar todos estos conceptos en el Trabajo.

La primera parte experimental del TFG consistió en el análisis de la herramienta de detección de atributos de Face++[1] con la base de datos LFW[2]. Para ello se tuvo que implementar un sistema que permitiese probar la base de datos a través del código muestra de la API de Face++. Una vez se probó esta herramienta con la base de datos LFW, se introdujo su primera variación, LFW frontal, parar ver los efectos que tenía este nuevo procesado de imagen en el detector de Face++.

Paralelamente, se realizó la segunda parte de este trabajo, que se trató de la introducción de un nuevo sistema de detección de Soft Biometrics proporcionado por la herramienta Microsoft Cognitive Services[3]. Para implementar esta herramienta se utilizó el lenguaje de Python. Los pasos seguidos para analizar esta herramienta fueron los mismos que en Face++. Primero se probó con la base de datos LFW y después con su variación frontal.

Una vez obtenidos los primeros resultados en cuanto a los rendimientos de los dos sistemas de estimación de Soft Biometrics se aumentó la robustez de estos resultados con la incorporación de dos bases de datos más. La base de datos CelebA[4] que contenía más de 200.000 fotos, y también se probó con una nueva variación de la base de datos LFW.

1. Se puede encontrar toda la información en https://www.faceplusplus.com
2. Se puede encontrar toda la información en http://vis-www.cs.umass.edu/lfw/
3. Se puede encontrar toda la información en https://azure.microsoft.com/es-es/services/cognitive-services/
4. Se puede encontrar toda la información en http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html

Para esta nueva variación de la base de datos LFW, se llevó a cabo un sistema de ecualización de los histogramas de la base de datos gracias a las herramientas de INface toolbox v2.0.

Una vez finalizado el estudio de los sistemas de estimación de Soft Biometrics, se aplicó la información y los resultados obtenidos en el desarrollo de un sistema que fusionaba una herramienta de reconocimiento de personas de Face++ junto con la información de los Soft Biometrics obtenida con el objetivo evaluar el efecto que ofrecería la información de los Soft Biometrics en el reconocimiento facial.

Por último, con todos los resultados obtenidos se realizó un análisis general y se escribió la memoria final del Trabajo, en la que se anotaban las conclusiones obtenidas junto con todos los experimentos realizados.

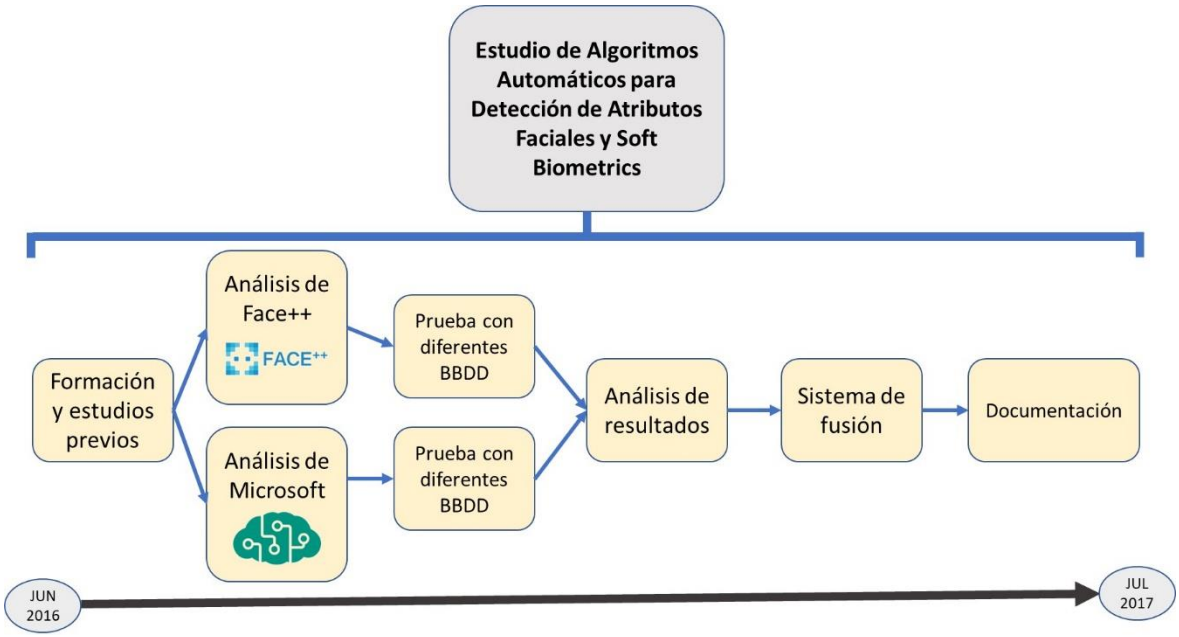


Figura 1-1: Diagrama de planificación y metodología seguidas

1.4 Organización de la memoria

La memoria de este trabajo consta de los siguientes capítulos:

- Capítulo 1: es la introducción al trabajo, en ella se explican las motivaciones y objetivos por los que se desarrolló este trabajo. También se comentará la metodología y la planificación de tiempos seguida.

- Capítulo 2: se definirá y se mostrará el estado del arte. En este capítulo se pretende introducir al lector los conceptos de biometría, reconocedor facial y Soft Biometrics. Por otro lado, se explicarán los trabajos previos en los que se ha basado y apoyado este TFG.

- Capítulo 3: en este tercer capítulo se explican las herramientas de detección de Soft Biometrics de Face++ y de Microsoft, explicando sus funcionamientos y restricciones. Estas son las dos herramientas estudiadas y analizadas durante este Trabajo.

- Capítulo 4: se mencionarán las dos bases de datos utilizadas: LFW y CelebA, se comentarán las variaciones realizadas sobre las imágenes de ellas y el análisis obtenido sobre sus etiquetados.

- Capítulo 5: se describen los sistemas de reconocimiento biométrico involucrados: sistema de reconocimiento facial de la herramienta Face++ y sistema basado en Soft Biometrics

- Capítulo 6: éste es el más extenso de los capítulos, puesto que en él se presentan y discuten todos los experimentos realizados, tanto en detección de Soft Biometrics como en fusión con el sistema de reconocimiento facial, y se muestran los resultados obtenidos de los que se sacaran unas breves conclusiones.

- Capítulo 7: en este último capítulo se expondrán las conclusiones obtenidas tanto a nivel local de cada parte como a nivel general. Por otro lado, se mencionará como afectará esta investigación a trabajos futuros.

2 Estado del arte

2.1 ¿Qué es el reconocimiento biométrico?

El reconocimiento biométrico consiste en la capacidad de detectar a un individuo a través del análisis de algunos de sus rasgos característicos, como son la cara, la firma manuscrita o las huellas dactilares. El uso de reconocedores biométricos es cada vez más común en todos los sistemas de seguridad.

Los sistemas biométricos funcionan registrando y comparando las características biométricas. Las características se registran según el tipo de reconocedor, desde imágenes para sistemas de reconocimiento facial o iris hasta ondas de voz para reconocedores de locutor. Por razones de eficiencia, en lugar de usar las características registradas directamente, es habitual extraer características de identificación de las muestras y codificar éstas facilitando el almacenamiento y la comparación. Aunque cada reconocedor biométrico es único y su implementación es complicada, la mayoría de ellos suelen seguir los dos mismos pasos:

- Fase de Registro: cuando un individuo usa por primera vez un sistema biométrico, sus características de identificación se registran como una referencia para la comparación futura. Esta referencia puede almacenarse en una base de datos central o en una tarjeta (o ambas) dependiendo de las necesidades de la aplicación.

- Fase de Reconocimiento: cuando se requiere reconocimiento biométrico, las características biométricas del individuo se adquieren nuevamente. Sin embargo, esta vez, las características de identificación son comparadas por el sistema con los modelos almacenados para determinar si existe una coincidencia cercana.

Existen dos modos para el reconocimiento biométrico: verificación e identificación. En la verificación, se reivindica una identidad y el proceso de comparación se limita a comprobar el modelo correspondiente a dicha identidad. En la identificación, ninguna reivindicación de identidad es necesaria y el sistema busca en su base de datos de modelos para encontrar si alguno de los modelos almacenados coincide con las características biométricas registradas.

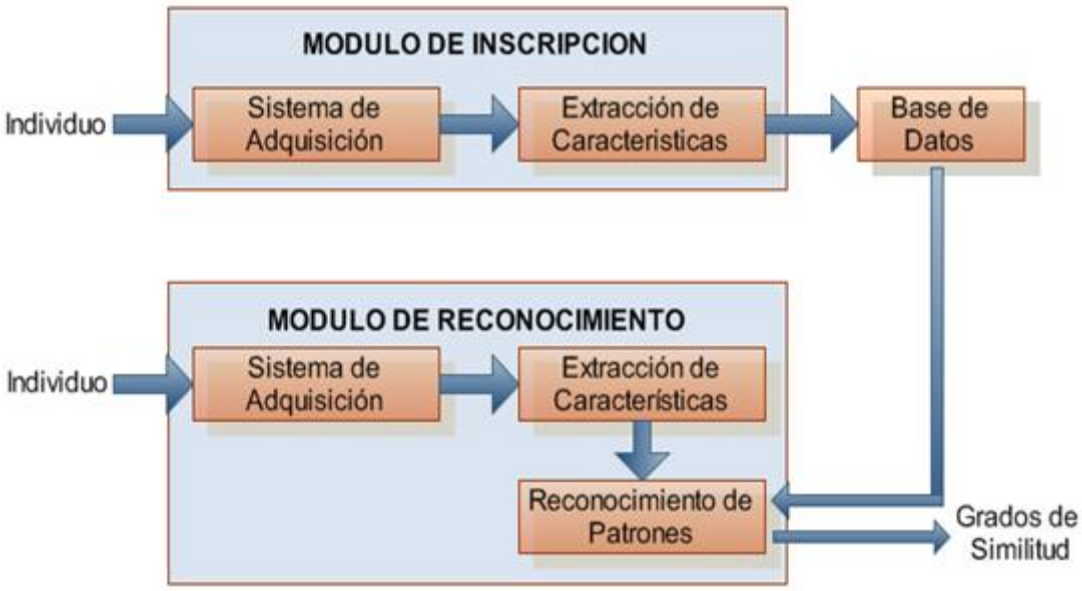


Figura 2-1: fases de un Reconocedor Biométrico. [1]

Tanto en el sistema de reconocimiento como en el sistema de detección hay que definir una medida de rendimiento, que vendrá determinada por el tipo de sistema que se quiera implementar. Ambos sistemas funcionan con un método de calcular puntuaciones, se comparan características y se devuelven puntuaciones indicando la similitud entre el conjunto de características comparadas. Pero a la hora de evaluar al sistema con el fin de determinar si un par de características pertenecen a la misma persona o no, es necesario establecer un umbral. En relación a este umbral se definen varias tasas: los falsos negativos y los falsos positivos, referidos al caso de decir que dos personas son la misma, cuando no lo son o al revés.

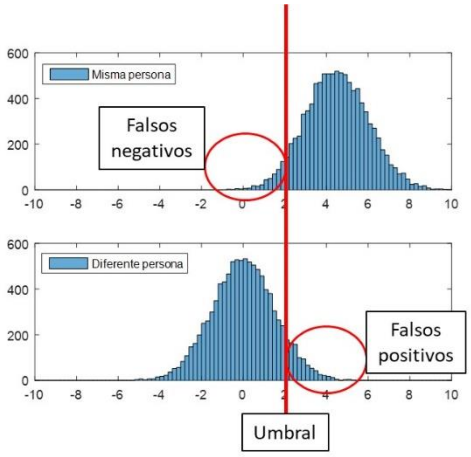


Figura 2-2: Histogramas de falsos positivos y negativos

La ubicación de este umbral de discriminación es una tarea bastante más complicada de lo que parece. Imaginemos un sistema de acceso a través de un reconocedor de firma manuscrita. El histograma superior de la *Figura 2-2* se corresponde a las puntuaciones obtenidas en la comparación de firmas de la misma persona, mientras que en el histograma

de abajo se muestran las puntuaciones obtenidas en la falsificación de esa firma. La recta roja delimita el umbral de discriminación. Y lo que queda en el interior de los círculos rojos son los falsos negativos (en el histograma superior) y falsos positivos (en el histograma inferior). Para este sistema tendríamos que barajar dos hipótesis:

- Si situamos el umbral de puntuación en niveles muy altos se minimiza la tasa de falsos positivos, aunque aumentaría la de falsos negativos. Es decir, lo más probable es que una persona acreditada no consiga entrar o bien tenga que repetir la prueba varias veces.
- Si situamos el umbral de puntuación en niveles muy bajos se minimiza la tasa de falsos positivos, pero aumenta la de falsos negativos. Es decir, un impostor podría falsificar la firma y acceder con ella fácilmente.

El punto en el que el número de falsos positivos es igual al número de falsos negativos se denomina Equal Error Rate (EER) [2]. Este es un valor muy utilizado en reconocimiento biométrico y cuanto más bajo sea significa que el sistema es más preciso.

Cabe destacar que los sistemas biométricos superan a los sistemas tradicionales en varios aspectos como la comodidad y facilidad de no tener que preocuparte de cualquier objeto externo o contraseña que se pueda extraviar u olvidar, dado que los sistemas biométricos trabajan con los rasgos característicos de cada uno.

Mientras que los rasgos biométricos se usan típicamente para reconocer individuos, es posible deducir otros tipos de atributos de esos mismos datos. Por ejemplo, a través de la información consistente facial se puede extraer información secundaria relativa a un individuo como puede ser la edad, el género, la etnia, el color de los ojos y otros atributos secundarios de la persona.

2.2 ¿Qué son los Soft Biometrics?

Soft Biometrics se define como cualquier característica del ser humano tanto física como dinámica que proporciona información relativa a un individuo, pero que por sí sola no permite el reconocimiento de esta persona. Un Soft Biometric debería ser una característica invariante en el tiempo, o al menos en un largo periodo de tiempo. [3]

Algunos ejemplos de Soft Biometrics son la etnia, la edad, la altura, el peso, el color del pelo, el color de los ojos, los tatuajes o las cicatrices. De ellos se puede observar como son rasgos propios de cada individuo, pero insuficientes para su identificación. Sabiendo que un individuo es negro y tiene los ojos azules se puede discriminar a gran parte de la población, pero con esos datos no podríamos identificar al individuo.

El número de Soft Biometrics aumenta a medida que va aumentado las tecnologías biométricas, cada uno de los Soft Biometric se puede clasificar como:

- Continuo o discreto, se puede distinguir si un Soft Biometric es continuo, como la edad, altura y peso que van definidos por un valor numérico, o discreto como la etnia o el género.
- Según la taxonomía: los Soft Biometrics se pueden clasificar como demográficos (edad, etnia o genero), geométricos (contorno facial o del cuerpo), médicos (peso

corporal, salud del corazón, índice de masa corporal) o materiales (gorras, gafas o ropa)

Como se observa en la *Figura 2-3*, según el tipo de reconocedor biométrico se usan unos Soft Biometrics u otros.

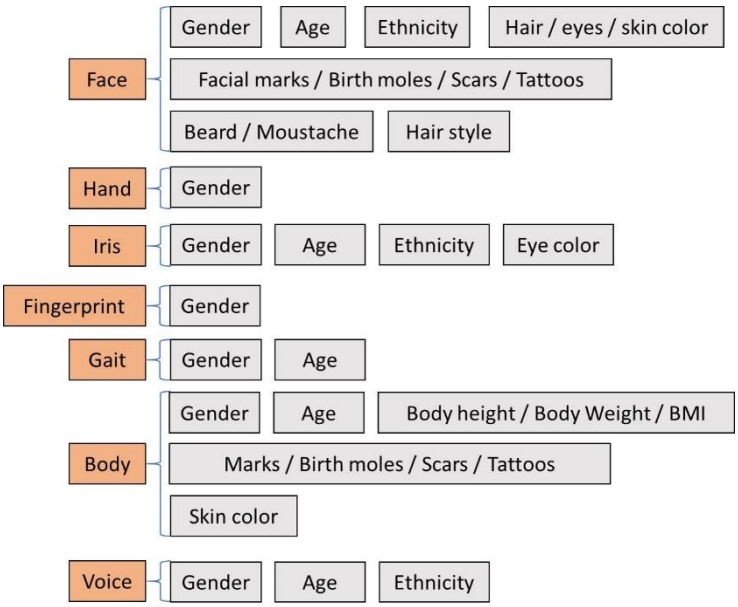


Figura 2-3: Soft Biometrics según el reconocedor biométrico

Por otro lado, no todos los rasgos pueden ser considerados Soft Biometric, para ello se necesitan una serie de requisitos:

- Un Soft Biometric tiene que ser un rasgo característico de cada individuo, pero tiene que aparecer en todos. Por ejemplo, la edad es característico de cada uno, pero todos tenemos una edad.
- Para ser considerado Soft Biometric tiene que perdurar en el tiempo. Por ejemplo, las gafas de ver podrían considerarse Soft Biometric al ser algo que una persona lleva de manera constante, mientras que, por el contrario, las gafas de sol no.
- Un Soft Biometric debe de ser discriminativo. Los hay menos discriminativos, como el género o más discriminativos con una cicatriz en la cara.
- El método de detección de un Soft Biometric no puede ser intrusivo, debe ser detectado sin la participación activa del usuario.

Es por tanto que se puede hacer una diferenciación entre lo que serían los Soft Biometrics (género, edad, altura, etnia) y lo que sería un atributo facial (gafas de sol, sonrisa).

2.3 Integración de Soft Biometrics para el reconocimiento

Las dos aportaciones principales de los Soft Biometrics en el reconocimiento de personas son:

- Reducir el espacio de búsqueda

- Fusionar con los rasgos característicos para obtener un mejor reconocimiento biométrico.

En este apartado lo que se pretende es explicar esta segunda aplicación. La posibilidad de implementar un sistema que gracias a la información obtenida de los Soft Biometrics y junto con el propio reconocedor biométrico ofrezca un mejor reconocimiento. [4]

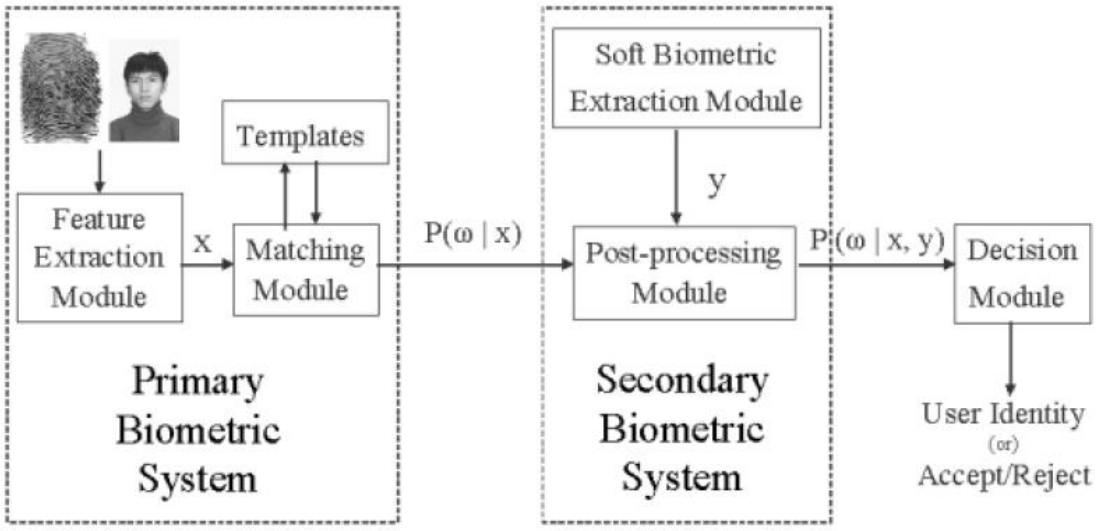


Figura 2-4: Integración de Soft Biometrics para el reconocimiento. [5]

Para la explicación de este sistema de fusión nos apoyaremos en la *Figura 2-4* correspondiente a un sistema de fusión de Soft Biometrics para el reconocimiento biométrico, en este caso de huella dactilar.

Imagine un sistema formado por una base de datos W = {w_1, w_2… w_n} donde cada una de las "w_i" con i = {1, 2… n} es un usuario. El primer paso del reconocedor es extraer los rasgos característicos del usuario a comparar, en este caso las huellas dactilares, y guardar estos datos en la variable "x". El segundo paso, como en todo reconocedor, es comparar esos datos con los de la base de datos en búsqueda de una coincidencia, siendo P(w_i|x) la probabilidad de que los datos "x" se correspondan con el usuario "w_i". Adicionalmente, al haberse añadido un sistema secundario de comparación de Soft Biometrics, esta probabilidad P(w_i|x), es de nuevo analizada por este segundo sistema formado por los datos Y = {y_1, y_2… y_m} correspondientes a los Soft Biometrics (género, color de piel…). [5]

La probabilidad de combinar estos dos vectores y determinar al usuario "w_i" se obtiene a través de la regla de Bayes:

$$P(wi \mid x, y) = \frac{p(y|wi)\, P\,(x|wi)}{\Sigma_{i=1}^{n}\, p(y|wi)\, P\,(x|wi)} \qquad (1)$$

Un problema de esta ecuacion, es que se da el mismo valor a cada Soft Biometric cuando en la realidad, algunos Soft Biometrics son más discriminativos que otros y por tanto más valiosos. Este problema se podría solucionar con un sistema de ponderación de Soft Biometrics.

Gracias a esta sistema, es posible mejorar el reconocimiento biométrico que un reconocedor podría ofrecer por si solo.

2.4 Trabajos previos

Para la realización de este TFG, se ha utilizado un etiquetado de la base de datos LFW creado por Beatriz Cid Fernández en su Trabajo de Fin de Grado titulado Interfaz Gráfica de Etiquetado de Atributos Faciales. [6]

En ese Trabajo se llevó a cabo un etiquetado de once rasgos faciales diferentes de cada una de las imágenes de la base de datos LFW. Para ello se creó una interfaz gráfica, como la de la *Figura 2-6*, implementada con el entorno de programación de Matlab que permitió un etiquetado rápido y sencillo de las 13233 fotos de la base de datos LFW.

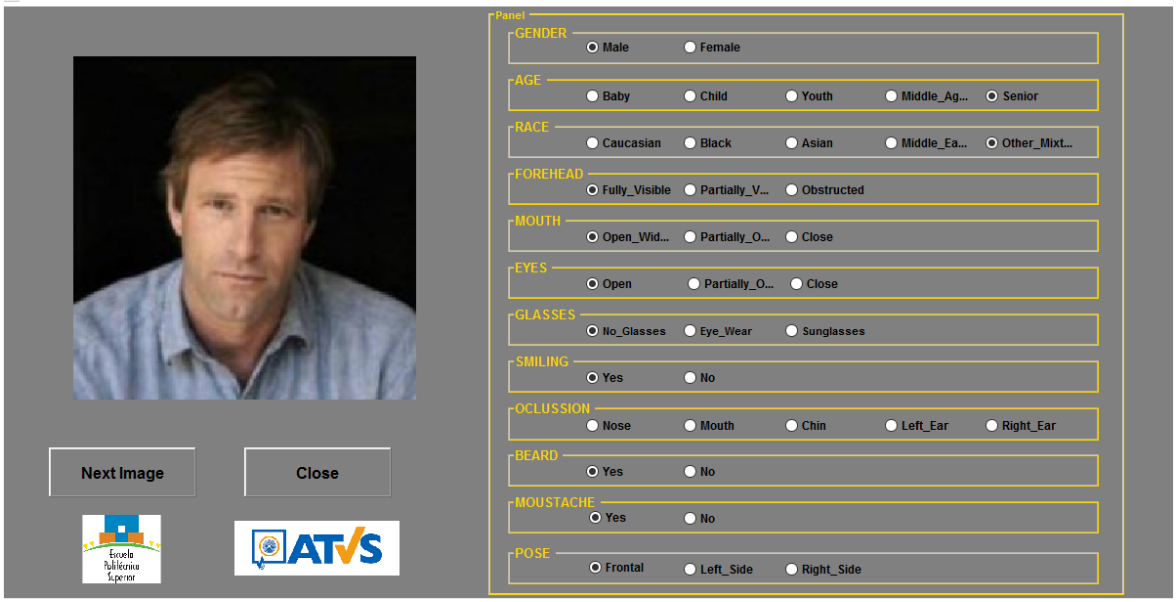


Figura 2-5: Ejemplo de la interfaz gráfica de etiquetado manual. [6]

Los once rasgos faciales analizados fueron: Gender, age, ethnicity, forehead, mouth, eyes, glasses, smiling, oclussion, beard y moustache.

3 Herramientas de detección Automática de Soft Biometrics

3.1 Face++

Face++ se trata de una herramienta que permite tanto la estimación de Soft Biometrics como realizar tareas de reconocimiento biométrico, verificación e identificación de personas en imágenes. Las dos herramientas principales usadas en este trabajo han sido la estimación de Soft Biometrics, utilizada para los experimentos de análisis de sistemas comerciales en detección de Soft Biometrics y la herramienta de reconocimiento utilizada para la fusión con los datos obtenidos de los Soft Biometrics.

El funcionamiento del sistema de estimación de Soft Biometrics consiste en analizar diferentes atributos característicos de un rostro mediante un sistema tecnológico de aprendizaje automático. Este sistema puede hallar los atributos o Soft Biometrics de edad, sexo, etnia, intensidad en la sonrisa, estado de los ojos, pose de la cabeza, calidad de la imagen de cara.

La herramienta de Face++ está basada en el avance del aprendizaje automático denominado *Deep Learning*. Este avance se implementa a través de redes neuronales (CNN) que procesan la información de cada imagen. Estas redes reciben la información de cada pixel de la imagen y mediante un sistema de arquitectura piramidal devuelve unos valores correspondientes a cada uno de los Soft Biometrics estimados. [7]

La estimación de Soft Biometrics de Face++ funciona únicamente con imágenes en los formatos JPG y PNG con un tamaño no superior a 2MB. Este sistema es capaz de detectar y analizar hasta cinco caras dentro de una misma imagen. De cada una de las caras que el sistema detecta se genera un identificador en la nube, que almacena y apunta a la información detectada con un tiempo de expiración de 72 horas.

Los parámetros que requiere la API de Face++ para la estimación de Soft Biometrics son unas claves de acceso que se obtienen registrándote en su página oficial, una imagen, como se ha mencionado antes, en formato JPG pero que esté codificada en binario, y dos valores boolenaos opcionales en función de los parámetros que se desean obtener, uno para los atributos faciales y el otro para las posiciones faciales de cada característica.

Un ejemplo de funcionamiento del sistema de estimación de Soft Biometrics sería el de la *Figura 3-1*, correspondiente a la actriz Emma Stone. Al introducir la imagen en el sistema de estimación de Face++ obtendríamos los siguientes resultados:

Figura 3-1: Ejemplo de funcionamiento del estimador de Face++. [8]

Como se puede observar con el resultado dado, en los atributos de género, gafas y etnia devuelve un valor de confianza con el que el sistema valora el resultado. En este ejemplo el detector de atributos estaría acertando en los atributos de género, gafa, edad y sonrisa, pero estaría fallando en el atributo de etnia del que con el valor de confidencia observamos que era en el que más dudaba.

3.2 Microsoft Cognitive

Microsoft Cognitive Services consiste en un conjunto de servicios, proporcionados a través de unas APIs, con los que a través del aprendizaje automático son capaces de proveer servicios de detección y compresión del habla, detección de emociones o el sistema de reconocimiento facial que se ha sido el que se ha usado en este proyecto.

La API de Face de Microsoft se trata de un servicio en la nube que proporciona los algoritmos de reconocimiento facial más avanzados. Dentro de todas las funciones que ofrece este sistema, sus dos principales son la estimación de Soft Biometrics y el reconocimiento de personas. Dado que durante este trabajo se ha estado desarrollando el primero de estos sistemas haremos énfasis en él.

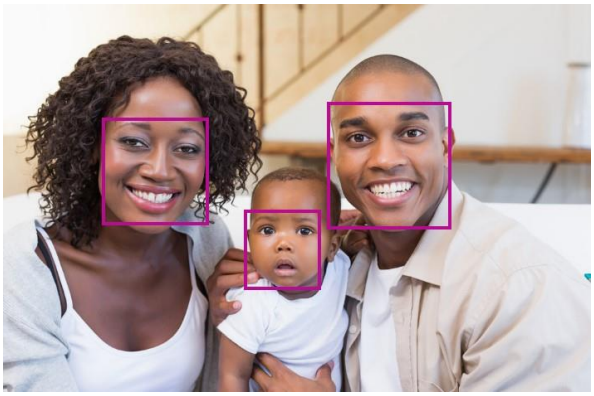


Figura 3-2: Ejemplo detección de caras de Microsoft. [9]

La herramienta de estimación de Soft Biometrics tiene la capacidad de detectar hasta 64 caras en una imagen. Además de detectar caras, esta herramienta proporciona la posibilidad de devolver también determinados atributos o Soft Biometrics de la cara. Estos son la edad, el género, la barba, la perilla, el bigote, la emoción, las gafas e la intensidad de la sonrisa.

El sistema de Microsoft funciona con imágenes en los formatos JPEG, PNG, GIF (el primer fotograma) y BMP, con un tamaño de la foto de entre 1 KB y 4MB. Para el desarrollo de

esta API Microsoft ofrece códigos ejemplo en los lenguajes de Curl, C, Java, PHP y Python, siendo este último lenguaje el utilizado para implementar el desarrollo.

De la misma manera que Face++, para el funcionamiento de la API de Microsoft, es necesario subscribirse al servicio de Microsoft Cognitive Services, el cual te dará la clave de acceso a la API. Junto con esta clave será necesario enviar una foto codificada en binario y dos variables opcionales correspondientes a si queremos los valores de los atributos y de las localizaciones de cada característica de la cara (landmarks).

A continuación, se muestra un ejemplo de cómo funciona el sistema de detección de atributos de Microsoft. La imagen 3-3, correspondiente a una niña con la cara pintada, codificada en binario y con formato JPG, es introducida en el sistema de estimación de Soft Biometrics. El estimador, tras comprobar las claves de la API, calcula los diferentes atributos y Soft Biometrics de la imagen y te devuelve una lista con todos ellos y sus intensidades como se muestra en esa misma imagen.

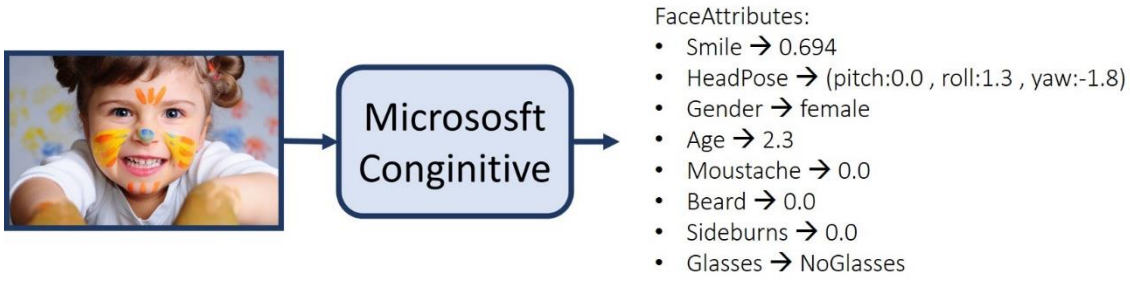


Figura 3-3: ejemplo de funcionamiento de Microsoft Cognitive

Como se puede observar, la herramienta de detección de atributos devuelve los atributos de barba, sonrisa, bigote y patillas con un valor de intensidad que oscila entre 0 y 1. Por otro lado, devuelve la edad con un decimal de precisión. Para el ejemplo de la imagen 3-3, el estimador de Microsoft estaría acertando en todos los Soft Biometrics, a excepción de la edad que, al ser desconocida la chica de la imagen, es imposible de corroborar.

Finalmente, y a modo de comparación en la tabla 3-1 se puede observar los atributos o Soft Biometrics que detecta Microsoft Cognitive frente a los que detecta Face++.

Tabla 3-1: comparación de atributos detectados por Face++ frente a Microsoft Cognitive

	Edad	Etnia	Género	Sonrisa	Barba	Gafas	Bigote
Face++	SI	SI	SI	SI	NO	SI	NO
Microsoft Cognitive	SI	NO	SI	SI	SI	SI	SI

4 Bases de datos

Para la realización de este Trabajo de Fin de Grado se han utilizado dos bases de datos diferentes que a continuación se mostrarán y analizarán:

4.1 Base de datos LFW

Es la primera de las bases de datos utilizada y en la que más nos hemos centrado durante todo el Trabajo. LFW (Labeled Faces in the Wild) se trata de una base de datos creada por The Computer Vision Laboratory en el Computer Scinece Department de la University of Massachusetts cuyo objetivo era crear una base de datos para mejorar el reconocimiento facial en entornos no controlados. El método por el cual se obtuvieron todas las imágenes de la base de datos LFW fue utilizando el sistema de detección de caras Viola-Jones llevándose a cabo en la biblioteca OpenCV. Algunas de las características principales de esta base de datos son:

- Se trata de una base de datos con 13233 fotos de personas obtenidas de la red. Cada foto está etiquetada con el nombre de la persona fotografiada junto con un identificador de tal manera que el nombre de cada una de las imágenes es único.
- En esta base de datos hay fotos de 5749 personas diferentes, de los cuales, 1680 individuos tienen dos o más imágenes diferentes en la base de datos. El resto de personas, 4069, tienen una única foto en la base de datos.
- Las imágenes de la base de datos son en su mayoría imágenes en color y vienen dadas en el formato JPEG de 250x250 pixeles. [10]

Gracias al etiquetado manual explicado en el punto 2.4, Trabajos previos, se ha obtenido un etiquetado de diversos atributos o Soft Biometrics de cada imagen, De los once atributos etiquetados se han seleccionado siete (Gender, Age, Ethnicity, Glasses, Smiling, Beard, Moustache) de los que se ha realizado un análisis previo a los experimentos para saber el número de fotos con cada atributo de la base de datos y su porcentaje correspondiente. Los datos obtenidos se enumeran en la siguiente tabla:

Tabla 4-1: Variabilidad de la base de datos LFW

Soft Biometric	Instancia	Número de fotos	%
Gender	Male	10260	77,53
	Female	2973	22,46
Age	Baby (<4 años)	11	0,08
	Child (5-17 años)	62	0,46
	Youth (18-39 años)	1708	12,90
	Middle Age (41-60 años)	8329	62,94
	Senior (>60 años)	3123	23,60

Ethnicity	White	10795	81,57
	Black	506	3,82
	Asian	733	5,53
	Indian	318	2,40
	Other Mixture	881	6,65
Glasses	No glasses	12445	94,04
	Eye wear glasses	344	2,60
	Sunglasses	444	3,35
Smiling	No	10573	79,89
	Yes	2476	18,71
Beard	No	12447	94,06
	Yes	786	5,94
Moustache	No	11908	89,98
	Yes	1325	10,01

Como se puede observar en el estudio, dentro de cada atributo los rasgos más predominantes son: Male en Gender, Middle Age en Age, White en Ethnicity, No Glasses en Glasses, Smile en Smiling, No Beard en Beard y No Moustache en Moustache.

Por otro lado, también es importante mencionar que el Computer Science Department desarrollo tres versiones de bases de datos LFW a partir de la base de datos con las imágenes originales, diferenciándose estas versiones por el tipo de alineación de sus imágenes. Quedando como resultado la base de datos LFW original, la base de datos LFW-funneled, la base de datos LFW-a y la base de datos Deep LFW. Este TFG se centra en la versión LFW funneled

En la figuro 4-1 se muestran ejemplos de imágenes de la base de datos LFW funneled.

Figura 4-1: Ejemplos de imágenes de la base de datos LFW funneled

Aparte de esos 3 tipos de bases de datos diferenciados según su alineación también se creó una última base de datos formada con las imágenes frontales. En la *Figura 4-2* se muestran ejemplos de imágenes de la base de LFW frontal:

Figura 4-2: Ejemplos de imágenes de la base de datos LFW frontal

Finalmente, se creó una última base de datos modificando la base de datos LFW funneled. Para la creación de esta base de datos se utilizó la herramienta INface toolbox v2.0, una herramienta de Matlab para la iluminación de imágenes en el reconocimiento de caras. De entre todas las funciones de esta herramienta se utilizó la función "Rank_Normalization" para la ecualización del histograma. [12]

El uso de esta herramienta se debe a que según varios estudios se ha demostrado que la ecualización del histograma de una imagen mejora significativamente el reconocimiento facial. En la imagen 4-3 se puede observar el procedimiento seguido para la ecualización y como se realizó la ecualización en cada uno de los canales RGB de manera independiente.

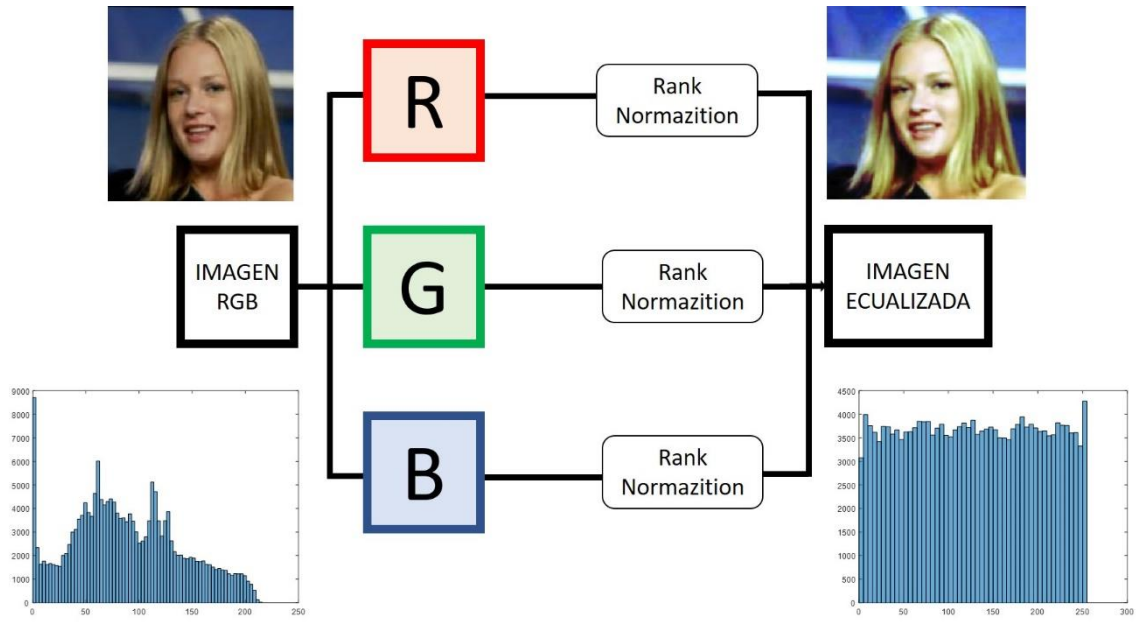


Figura 4-3: Procedimiento de ecualización del histograma

Algunos otros ejemplos de imágenes de la base de datos LFW funneled tras la ecualización del histograma se pueden ver en la *Figura 4-4*:

Figura 4-4: Ejemplos de imágenes de la base de datos LFW tras la ecualización del histograma

4.2 Base de datos CelebA

Se trata de una base de datos con 202599 fotos de personas obtenidas de la red. La base de datos fue creada por el laboratorio multimedia de la Chinese Univerity of Hong Kong.
Cada una de las más de doscientas mil imágenes esta etiquetada con cuarenta atributos diferentes. La base de datos de CelebA está formada por identidades de 10177 personas diferentes y el objetivo principal por el que fue creada esta base de datos fue el testeo de sistemas de estimación de atributos y de reconocimiento de personas.

Algunos ejemplos de imágenes de la base de datos son:

Figura 4-5: ejemplos de imágenes de CelebA. [13]

De entre los cuarenta atributos que nos proporciona CelebA se ha llevado a cabo un análisis de siete atributos, escogidos en función de la importancia que se le dará durante el desarrollo del proyecto, los resultados obtenidos en cuanto al número de imágenes de cada atributo en la base de datos han sido los siguientes:

Tabla 4-2: Variabilidad de la base de datos CelebA

Soft Biometric	Instancia	Número de fotos	%
Gender	Male	84434	41,67
	Female	118165	58,32
Age	Young	156734	77,36
	Others ages	45865	22,63
Glasses	No glasses	189406	93,48
	Eye wear glass	13193	6,51
Smiling	No	104930	51,79
	Yes	97669	48,20
Beard	No	169158	83,49
	Yes	33441	16,50
Moustache	No	194182	95,84
	Yes	8417	4,15
Sideburn	No	191150	94,34
	Yes	11449	5,65

Como se puede observar a partir de los resultados obtenidos los atributos que más predominan en la base de datos son los de Male para Gender, Young para Age, no glasses, no smile, no beard, no moustache y no sideburns. Cabe destacar que, en el aspecto de la edad, CelebA, solo hace discriminación entre joven u otra edad.

5 Sistemas de Reconocimiento

5.1 Facial

El sistema de reconocimiento facial está basado en la herramienta de comparación de caras de Face++, este sistema funciona obteniendo una puntuación de confianza que vendrá definida por la probabilidad de que dos imágenes de caras pertenezcan a la misma persona. Se trata de una herramienta que calcula la similitud entre caras.

Este sistema de Face++ es muy útil en la verificación de usuarios mediante la cara y en la identificación de personas. Al igual que en la detección de atributos, su funcionamiento está basado en el *Deep Learning* y las redes neuronales. Cabe destacar que esta herramienta se encuentra presente en empresas financieras como sistema de seguridad para el pago en línea.

Un ejemplo de funcionamiento de este sistema sería el de la comparación entre las imágenes del actor Brad Pitt con dos looks completamente diferentes. El resultado de esta comparación sería el siguiente:

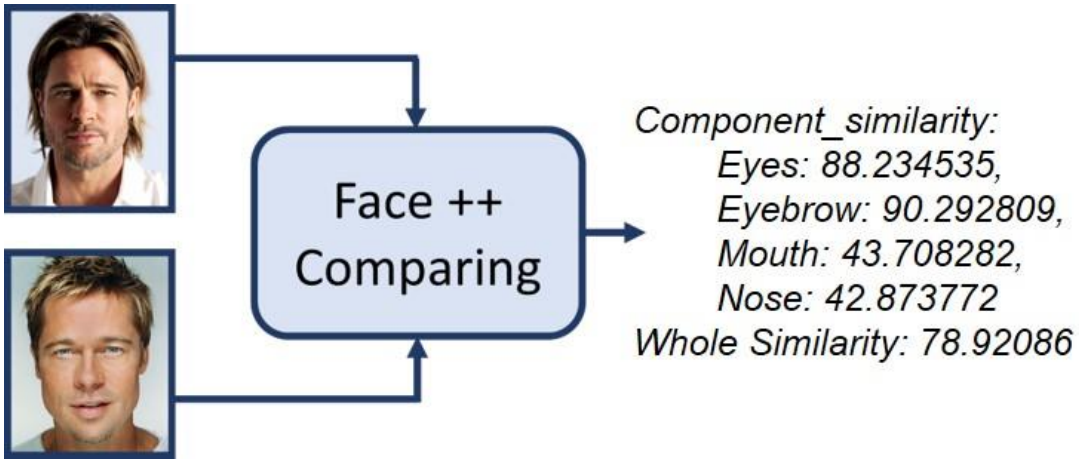


Figura 5-1: ejemplo de reconocimiento de personas en Face++

Como se puede observar en la *Figura 5-1*, la herramienta de comparación devuelve diferentes tipos de puntuaciones, todas entre uno y cien. Las definidas como componentes de similitud, que se corresponden a la similitud en las diferentes zonas. Luego se devuelve una puntuación global. Para este ejemplo, se puede ver como a pesar de que las apariencias del actor en cada foto son diferentes, el sistema devuelve una puntuación de similitud de 88 sobre 100, siendo esta bastante alta.

5.2 Soft Biometrics

El sistema de reconocimiento de Soft Biometrics es también un sistema de comparación de imágenes, solo que, en este caso, las características discriminatorias son el conjunto de Soft Biometrics.

Para explicar el sistema de comparación de Soft Biometrics es necesario entender antes los significados de distancia Hamming y distancia Euclídea, ya que son estas dos distancias, las que delimitaran la similitud entre dos Soft Biometrics:

- Distancia Hamming: la distancia Hamming mide el número de bits que difieren dos números binarios, por ejemplo:

1	0	0	1	0	1	1	0
1	0	1	1	0	0	1	1

Figura 5-2: ejemplo de cálculo de la distancia Hamming, siendo esta tres.

- Distancia Euclídea: es la distancia más pequeña que une dos puntos en un espacio euclídeo, esta distancia se obtiene gracias al teorema de Pitágoras. Si tenemos dos puntos P (P1, P2) y Q (Q1, Q2) en un sistema bidimensional la distancia entre ellos se obtendría mediante:

$$Distancia = \sqrt{(P1 - Q1)^2 + (P2 - Q2)^2}$$

Figura 5-3: ejemplo de cálculo de distancia Euclídea en dos dimensiones

Atendiendo a la naturaleza de cada Soft Biometric se decide utilizar un tipo de distancia u otro. Concretamente, la edad es el único Soft Biometric que se ha comparado utilizando la distancia Hamming. La razón de esta diferenciación es porque la edad es un Soft Biometrics que sí guarda una relación proporcional con su instancia, es decir, confundir un valor de etiqueta cero en edad con un valor tres equivaldría a confundir un bebé con un hombre de mediana edad siendo este un error más grave que confundir uno de mediana edad con un adulto. A su vez, si por ejemplo aplicamos esos mismos valores de confusión en el Soft Biometric de la etnia equivaldría a confundir un hombre de etnia blanca con un indio siendo este un error no proporcional.

6 Integración, pruebas y resultados

6.1 Pruebas con Face++

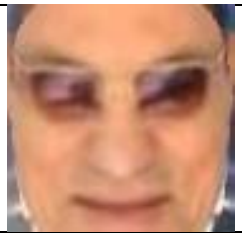
Para trabajar con el reconocedor de Face++ se utilizó la herramienta Matlab dado que proporcionaban la API con este lenguaje. El proceso que se realizó consistió en que, a cada una de las imágenes de las cuatro variedades de bases de datos, LFW funneled, LFW frontal, LFW ecualizados los histogramas y CelebA, se les aplicó el detector de atributos desarrollado por Face++, por el cual de cada imagen se obtuvo un conjunto de datos con los diferentes resultados de los Soft Biometrics. Una vez obtenidos esos datos de cada una de las imágenes comenzó el análisis del sistema.

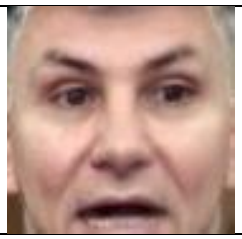
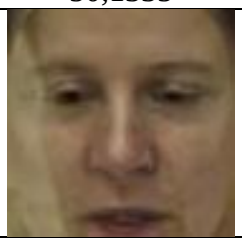
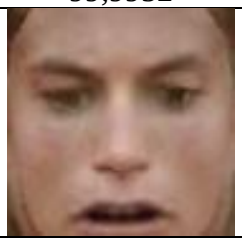
Los dos primeros aspectos que se tuvieron en cuenta antes de analizar el rendimiento en el acierto de los atributos fueron contabilizar el número de imágenes en las que el detector de Face++ no detectaba ninguna cara, guardando este número para su posterior análisis, y ver que estaba ocurriendo con las imágenes en las que Face++ detectaba varias caras, en las que se llegó a la conclusión de que las ordenaba según el tamaño de la cara, pero que debido a las irregularidades de las imágenes en las diferentes bases de datos, estos resultados podrían provocar errores en el análisis del rendimiento del sistema por lo que fueron guardados y apartados.

Una vez filtrados los resultados y apartados los casos en los que no se detectaba ninguna cara o bien había múltiple detección se llevó a cabo el análisis propio de los resultados obtenidos del detector en cuanto a los atributos de edad, género, gafas, etnia y sonrisa. De dónde se obtuvieron resultados en cuanto al rendimiento del sistema, expuestos en el punto 6.3 y resultados propios de cada uno de los atributos debido a sus valores de confianza e intensidad, que están explicados a continuación.

Del Soft Biometric del género cabe destacar que las imágenes en las que el detector de Face++ falló con un mayor valor de confianza son:

Tabla 6-1: errores en la estimación de género de Face++

Base de datos	LFW funneled	LFW frontal	LFW ecualizada	CelebA
Predice mujer con mayor confianza				
Confianza	99.9858	99,9942	99,9892	99,9997

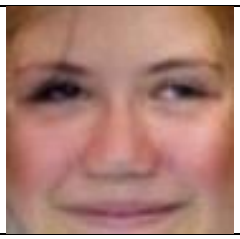
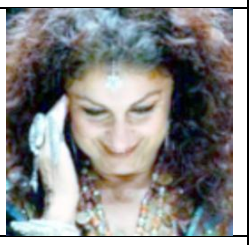
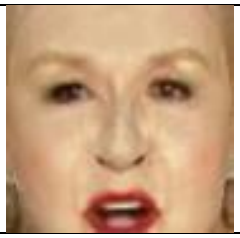
Predice mujer con menor confianza				
Confianza	50,0901	50,1335	50,0776	50,0345
Predice hombre con mayor confianza				
Confianza	99,9459	99,9952	99,9746	99,9999
Predice hombre con mayor confianza				
Confianza	50,0872	50,0192	50,1376	50,0088

Como se puede observar la mayoría de los errores podrían haber sido cometidos por el ser humano distinguiendo manualmente entre ambos géneros. Los hombres con el pelo largo causan bastante confusión al reconocedor. Por otro lado, en las imágenes frontales los casos a estimar en el género ofrecen bastantes dudas.

En cuanto al Soft Biometric de la edad, que es el Soft Biometric con más errores en las estimaciones, podemos observar algunos de sus errores en la *Tabla 6-2*:

Tabla 6-2: errores en la estimación de edad de Face++

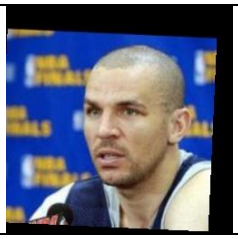
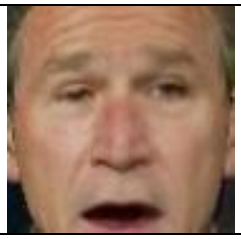
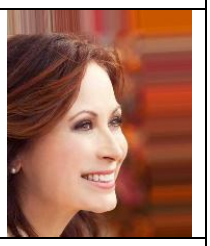
Base de datos	LFW funneled	LFW frontal	LFW ecualizada
Errores en la detección de bebés	Sin errores		
Edad estimada	Sin errores	17	5

Errores en la detección de niños			
Edad estimada	33	33	48
Errores en la detección de jóvenes			
Edad estimada	60	54	61
Errores en la detección de mediana edad			
Edad estimada	2	2	2
Errores en la detección de senior			
Edad estimada	9	3	4

En primer lugar, hay que mencionar que la base de datos CelebA se descartó en la detección de edad, como se explica en la sección 6.3, y por tanto la comparación se realizó entre las otras tres bases de datos. En la detección de bebés sorprende el caso de la imagen frontal en la que a un bebé de 3 años se le asigna una edad de 17. El error que se comete en el niño de LFW funneled parece estar debido a que el gesto que realiza le da una apariencia mayor. Cabe destacar el caso de mediana edad con mayor error, en el que tanto en la base de datos LFW funneled como en su variación con la ecualización de histogramas, una mujer, aparentemente mayor es detectada como un bebé de dos años. A pesar de que la señora no está mirando a la cara, el error parece claro.

Para la detección de sonrisa, se tuvo que realizar una discriminación de un umbral de intensidad en el que se consideraba que la persona de la imagen estaba sonriendo. La intensidad de sonrisa venia definida entre 0 y 100 y se consideró, tras la realización de diferentes pruebas, que el mejor umbral estaba situado en el 46.

Tabla 6-3: errores en la estimación de sonrisa de Face++

Base de datos	LFW funneled	LFW frontal	LFW ecualizada	CelebA
Debería detectar sonrisa y falla				
Intensidad (0-100)	0,4536	0,5543	0,5804	0,2297
Debería detectar sin sonrisa y detecta con				
Intensidad (0-100)	98,6106	99,1138	97,3597	99,5310

Como se puede observar en los errores en los que debería detectar sin sonrisa y detecta con sonrisa, el error se puede deber al aspecto alegre de la persona en la imagen, que aun sin llegar a sonreír sí que muestra un aspecto feliz. En cuanto a los fallos en la detección de sonrisa, Face++, no detecta sonrisa en imágenes que tienen la boca abierta y están mostrando los dientes, pero sin llegar a sonreír. Un error con menor explicación es el fallo en la imagen de CelebA, que podría deberse a que la imagen esta de lado, pero que aun así la sonrisa es clara. No obstante, otro posible motivo de los errores en la detección de sonrisa sería la discrepancia entre el criterio decidido por Face++ para establecer que se está sonriendo y el criterio elegido en el etiquetado manual.

El atributo o en ocasiones Soft Biometric de las gafas no parece dar muchos problemas a la hora de ser detectado, y es que, dada la oscuridad de unas gafas de sol, la discriminación entre ellas y las gafas de ver parece relativamente sencilla. Igualmente, el detector de atributos de Face++ ha cometido algún error notable que a continuación se muestra:

Tabla 6-4: errores en la estimación de gafas de Face++

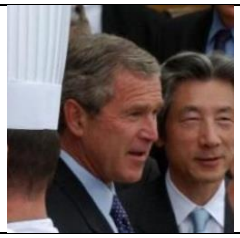
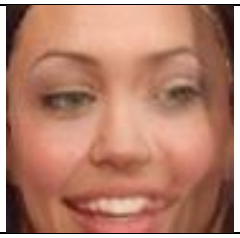
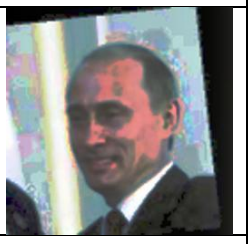
Base de datos	LFW funneled	LFW frontal	LFW ecualizada
Debería detectar sin gafas, pero detecta algún tipo de gafas			
Confianza	99.4232	99.3612	99.1960

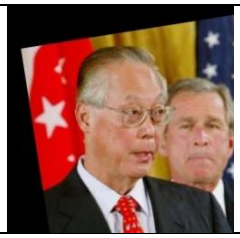
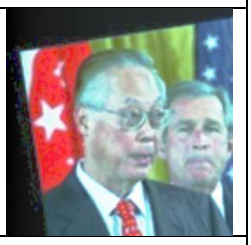
No detecta gafas de ver con mayor confianza			
Confianza	99,7044	99,9427	99.8747
Detecta gafas de sol con mayor confianza			
Confianza	99.9706	99.9919	99,9624

Como se puede observar los errores en los que detecta gafas y no deberían detectarse son bastante comprensibles dado que hasta un ser humano etiquetando manualmente se tendría que acercar mucho en la imagen para poder decidir. En cuanto a los fallos de detección de gafas de ver, se observa que los casos en los que se tratan de gafas de ver sin borde y completamente transparentes son los que más complican al sistema. En el último de estos errores, sí que se aprecian más claramente las gafas, pero al estar debajo de la altura de los ojos salen de lo que sería la zona de mayor análisis en la detección de gafas y dificulta su detección. Es curioso, que, en la misma imagen en sus tres variaciones, Face++ siempre detecta gafas de sol, pudiendo deberse este error a las ojeras de la persona de la imagen.

Por último, queda la etnia, quizás el atributo más difícil de clasificar debido a la cantidad de factores que intervienen en él. A continuación, se muestran algunos de los mayores errores cometidos detectando etnia.

Tabla 6-5: errores en la estimación de etnia de Face++

Base de datos	LFW funneled	LFW frontal	LFW ecualizada
Debería detectar blanco y detecta otra etnia			
Confianza	99.9688	99.9795	99.9727
Etnia detectada	Asiático	Asiático	Negro

Debería detectar negro y detecta otra etnia			
Confianza	99.3236	99.5327	99.2162
Etnia detectada	Blanco	Blanco	Blanco
Debería detectar asiático y detecta otra etnia			
Confianza	99.9764	99.9961	99.9932
Etnia detectada	Blanco	Blanco	Negro

En este último análisis no se ha tenido en cuenta la base de datos CelebA dado que en su etiquetado no proporcionaba información acerca de la etnia de la persona en cada imagen. En cuanto a los resultados, destacan sobre todo los fallos en las fotos detectando gente negra, dado que en los fallos de la imagen de LFW funneled y LFW frontal se observa claramente la piel oscura de la persona.

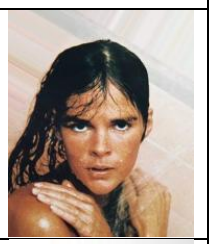
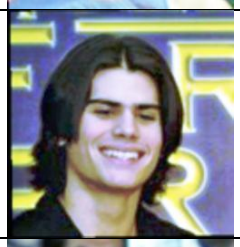
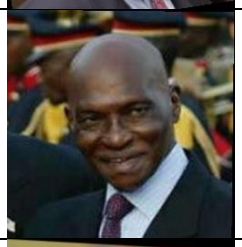
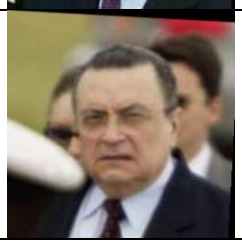
6.2 Pruebas con Microsoft

Por otro lado, el análisis de sistema de estimación de Soft Biometrics de Microsoft se realizó inicialmente en el lenguaje de Python y después con los resultados obtenidos se trabajó en Matlab

Lo primero que se hizo fue, a través del código de muestra que ofrecía Microsoft, se implementó un sistema que iba recorriendo cada una de las fotos de las diferentes bases de datos para guardar sus valores en forma binaria, con estos datos y la claves específicas se hacía una llamada a la API y está devolvía un resultado con los datos de los atributos detectados, como se explicó en el apartado 3.2, cada uno de estos resultados fue guardado en una lista almacenada para poder trabajar con los resultados en Matlab.

El método utilizado en Matlab para el análisis de los resultados obtenido fue muy similar al utilizado para los resultados de Face++, diferenciándose sobre todo en los atributos analizados, dado que Microsoft y Face++ detectan algunos atributos diferentes. Y la otra gran diferencia fue la forma de analizar los resultados porque la anotación de cada uno de los dos sistemas era diferente. Dado que el sistema de Microsoft no devolvía valores de confianza con la detección de los atributos no haremos un inciso individual en sus resultados. No obstante, en la *Tabla 6-6* se pueden ver ejemplos de algunas imágenes en las que se detectan mal los Soft Biometrics con el estimador de Microsoft.

Tabla 6-6: errores en la estimación de Soft Biometrics con Microsoft Cognitive

Base de datos	LFW funneled	LFW frontal	LFW ecualizada	CelebA
Detecta hombre y debería detectar mujer				
Detecta barba y debería detectar sin barba				
Detecta sin sonreír y debería detectar sonriendo				
Detecta sin gafas y debería detectar con gafas				

Se observa que se tratan de errores graves en su mayoría, aunque al no devolver Microsoft un valor de confianza no se puede saber hasta qué punto el estimador de Microsoft estaba dudando.

En el sistema de detección de atributos de Microsoft, el atributo de la sonrisa venía delimitado por un número del cero al uno, realizando un barrido de comprobaciones en el rendimiento se llegó a la conclusión de que el umbral que mejor delimitaba si en la imagen se estaba sonriendo debía estar en el cero, quedando etiquetada cualquier foto con intensidad de sonrisa mayor que cero como sonriendo.

Por último, a diferencia de la herramienta de Face++, Microsoft detecta el atributo de la barba. Y es por tanto que expondremos los resultados a continuación, en la *Tabla 6-7*, al no tener herramienta con la que compararse.

Tabla 6-7: Tasas de acierto y fallo estimación de barba con Microsoft

	Beard	Nº	%
LFW funneled	Tasa de acierto	9584	76,86

	Tasa de error	2885	23,13
LFW frontal	Tasa de acierto	8548	67,30
	Tasa de error	4152	32,69
LFW ecualizada	Tasa de acierto	8821	70,76
	Tasa de error	3645	29,23
CelebA	Tasa de acierto	108083	55,53
	Tasa de error	86544	44,46

Como se puede observar, la estimación de la barba con la herramienta de Microsoft Cognitive es relativamente buena, dando tasas superiores al 70% en las bases de datos LFW menos en la base de datos LFW frontal, donde posiblemente la deformación en las caras de las imágenes sea la causa del aumento de errores. Por otro lado, la peor tasa de acierto se da con la base de datos CelebA, probablemente debido a discrepancias en el etiquetado.

6.3 Comparación entre Face++ y Microsoft

Tanto el detector de atributos de Face++ como el de Microsoft usan métodos similares, pero no iguales, son esas diferencias las que provocan que uno de los dos detectores sea más potente que el otro.

6.3.1 Tiempo computacional

En cuanto a la potencia de cada uno de los detectores, lo primero que se midió fue la velocidad del sistema, para ello se realizó un experimento seleccionando las cien primeras fotos de la base de datos CelebA. A cada uno de los dos detectores se les paso estas fotos para que detectaran los atributos de cada una de ellas y los resultados obtenidos en cuanto a los tiempos de detección fueron los siguientes:

Tabla 6-8: Comparación entre tiempos computacionales

	Microsoft	Face++
Tiempo total dedicado para la detección de los atributos de las 100 fotos	108.81 segundos	2229.19 segundas
Tiempo medio dedicado a cada foto	1.09 segundos	22.29 segundos
Máximo tiempo dedicado a una foto	1.57 segundos	6.58 segundos
Mínimo tiempo dedicado a una foto	0.51 segundos	91.20 segundos

Dado que para ambas detecciones se utilizó la misma conexión de 300Mb de ancho de banda y se realizó el experimento en el mismo horario, se puede observar como la velocidad del detector de Microsoft es mucho menor que la del de Face++, siendo la estimación de Microsoft casi doscientas veces más rápida.

6.3.2 Rendimiento y precisión

Una vez analizados los resultados en cuando a la velocidad de cada herramienta se pasó a estudiar el rendimiento y la precisión de cada sistema para cada una de las bases de datos utilizadas durante este proyecto. En la *Tabla 6-9* se puede observar los rendimientos del sistema en cuanto a detección de una cara en la imagen, de múltiples caras o de ninguna cara, siendo esto último el dato más negativo de cada herramienta

Tabla 6-9: Comparación entre tasa de acierto y fallo en detección de caras

	GENERAL	Face++		Microsoft	
		Nº	%	Nº	%
LFW funneled	No detecta la cara	225	1,70	23	0,17
	Múltiple cara	488	3,68	741	5,60
LFW frontal	No detecta la cara	1778	13,43	533	4,02
	Múltiple cara	0	0	0	0
LFW ecualizada	No detecta la cara	328	2,47	45	0,34
	Múltiple cara	442	3,34	722	5,45
CelebA	No detecta la cara	13041	6,43	1542	0,76
	Múltiple cara	2194	1,08	6426	3,17

Como se puede observar de la tabla anterior, el detector de caras de Microsoft es mejor a la hora de detectar caras en una imagen, por consiguiente, su detección de múltiples caras en una imagen es mayor. Otro factor importante que afecta en la detección de múltiples caras es que Microsoft puede detectar hasta 40 caras en una imagen mientras que Face++ tiene puesto su máximo en 5 caras. De todos modos, se observa como ambos sistemas dan muy buenos resultados en la detección de caras detectando ambos más del 90% de las caras. La base de datos en la que ambos sistemas más fallan es la base de datos LFW frontal en la que las imágenes están más deformadas.

Una vez hemos obtenido los rendimientos en cuanto a la detección de caras de cada uno de los sistemas pasaremos a evaluarlos y compáralos en la detección de cada Soft Biometric y atributo de manera individual. El primero de los Soft Biometrics estudiados será el género.

Tabla 6-10: Comparación entre tasa de acierto y fallo en estimación de género

	GÉNERO	Face++		Microsoft	
		Nº	%	Nº	%
LFW funneled	Tasa de acierto	12058	96,30	12302	98,66
	Tasa de error	462	3,69	166	1,33
LFW frontal	Tasa de acierto	10553	92,12	12308	96,92
	Tasa de error	902	7,87	391	3,078
LFW ecualizada	Tasa de acierto	11861	95,16	12251	98,28
	Tasa de error	602	4,83	214	1,71
CelebA	Tasa de acierto	180240	96,19	191835	98,56
	Tasa de error	7124	3,80	2792	1,43

En la detección de género el detector de Microsoft sigue siendo mejor frente al detector de Face++ superándolo en casi un 2 % tanto en LFW como en CelebA. Aunque cabe destacar que ambos sistemas detectan el género con una tasa de acierto superior al 95%. De nuevo la base de datos LFW frontal es la que da peores resultados, probablemente debido a la deformación que sufrieron las caras por la alineación frontal, y la base de datos LFW ecualizada se mantiene con resultados muy parecidos a los de las otras bases de datos, aunque ligeramente peores.

El segundo atributo analizado fue el de la edad. Para el análisis de la edad no se tuvo en cuenta la base de datos CelebA dado que, aunque sí que había un etiquetado booleano de si la persona de la imagen era joven o no, este dato era demasiado confuso y provoca demasiados errores al no estar delimitado cual era la franja de edad que CelebA consideraba joven, puesto que de las 202599, más del setenta por ciento correspondían a personas jóvenes, siendo una cifra extraña al ser considerado franja joven entre los 18 y 39 años.

Tabla 6-11: Comparación entre tasa de acierto y fallo en estimación de edad

	EDAD	Face++		Microsoft	
		Nº	%	Nº	%
LFW funneled	Tasa de acierto	5140	41,05	7844	62,91
	Tasa de error	7380	58,94	4624	37,08
LFW frontal	Tasa de acierto	4230	36,92	7531	59,30

	Tasa de error	7225	63,07	5168	40,69
LFW ecualizada	Tasa de acierto	4201	33,70	7725	61,97
	Tasa de error	8262	66,29	4740	38,02

La edad es sin duda el Soft Biometric en el que más fallan ambos detectores. El detector de Microsoft tiene una tasa de acierto del 60 %, que se aleja bastante de sus tasas de acierto en otros Soft Biometrics. Mientras que el detector de Face++ tiene una tasa de acierto muy baja que no supera el 40%, es decir, Face++ falla más que acierta detectando la edad. Aspectos como que no todo el mundo aparenta la edad que tiene o que para detectar la edad se tienen en cuenta muchos aspectos son los que provocan que este atributo es el más complicado de detectar.

Por otro lado, tenemos el Soft Biometric de la etnia, que a priori al igual que la edad parece otro atributo bastante difícil de detectar. Este Soft Biometric no es detectado por la herramienta de Microsoft por lo que en la *Tabla 6-12* solo figuran los resultados de la detección de Face++.

Tabla 6-12: Tasas de acierto y fallo estimación de etnia con Face++

	ETNIA	Face++	
		Nº	%
LFW funneled	Tasa de acierto	10522	84,04
	Tasa de error	1998	15,95
LFW frontal	Tasa de acierto	9670	84,41
	Tasa de error	1785	15,58
LFW ecualizada	Tasa de acierto	10478	84,07
	Tasa de error	1985	15,92
CelebA	Tasa de acierto	NADA	NADA
	Tasa de error	NADA	NADA

Como se puede observar, Face++ detecta la etnia con un 84% de acierto siendo este un dato bastante alto. Uno de los resultados más sorprendentes en la detección de etnia es que las bases de datos LFW frontal y LFW ecualizada ofrecen resultados ligeramente superiores a los de LFW funneled. Esto se puede deber a que en esas bases de datos los tonos de piel se ven significativamente mejor siendo este atributo clave en la detección de etnia.

El siguiente atributo, o en ocasiones Soft Biometric, del que se han investigado los resultados es el de las gafas en las imágenes. La situación de que la persona de la imagen lleve gafas de ver, gafas de sol o no lleve gafas. Los resultados de esta detección se pueden ver en la *Tabla 6.13*.

Tabla 6-13: Comparación entre tasa de acierto y fallo en estimación de gafas

	GAFAS	Face++		Microsoft	
		Nº	%	Nº	%
LFW funneled	Tasa de acierto	12203	97,469	12142	97,3853
	Tasa de error	317	2,531	326	2,6146
LFW frontal	Tasa de acierto	11019	96,193	12057	94,944
	Tasa de error	436	3,806	642	5,055
LFW ecualizada	Tasa de acierto	11888	95,386	11671	93,630
	Tasa de error	575	4,613	794	6,369
CelebA	Tasa de acierto	181310	96,768	189568	97,400
	Tasa de error	6054	3,231	5059	2,599

Aunque la diferencia sea mínima, en esta detección la herramienta de Face++ supera a la de Microsoft en todas las bases de datos menos en CelebA. Se trata de uno de los pocos atributos en los que Face++ supera a Microsoft, y es un atributo en el que ambos obtienen muy buenos resultados con un índice de acierto de más del 95%.

El último de los atributos en los que se puede hacer comparación entre Face++ y Microsoft es el de la sonrisa. En la detección de sonrisa, como se explicó en el punto 4, se ha llevado a cabo una discriminación de la puntación situando un umbral para delimitar cuando están sonriendo o no. En Face++ que devuelve la intensidad de sonrisa entre 0 y 100 ese umbral se situó en 46, mientras que en Microsoft que devuelve la intensidad entre 0 y 1 ese umbra se ubicó en el 0. Los resultados obtenidos se pueden ver en la *Tabla 6-14*.

Tabla 6-14: Comparación entre tasa de acierto y fallo en estimación de sonrisa

	SONRISA	Face++		Microsoft	
		Nº	%	Nº	%
LFW funneled	Tasa de acierto	11091	88,586	5324	42,701
	Tasa de error	1429	11,413	7144	57,298
LFW frontal	Tasa de acierto	10040	87,647	5324	42,701
	Tasa de error	1415	12,352	7144	57,298

LFW ecualizada	Tasa de acierto	10994	88,213	4985	39,991
	Tasa de error	1469	11,786	7480	60,008
CelebA	Tasa de acierto	166456	88,840	103082	52,963
	Tasa de error	20908	11,159	91545	47,036

En los resultados de la *Tabla 6-14*, se aprecia como el detector Face++ detecta mucho mejor las sonrisas que el detector de Microsoft. Este es el atributo mejor detecta Face++ frente a Microsoft, en el que con una tasa de acierto de más del 88% casi logra doblar a Microsoft que se queda en tasas por debajo del 45% de acierto. Se puede pensar que la mala detección de Microsoft viene determinada por la ubicación del umbral de sonrisas, pero se realizó un barrido buscando el umbral que diera mejores resultados y por ello se estableció ese.

6.4 Pruebas con fusión

El sistema de fusión de un reconocedor facial con el sistema de reconocimiento de Soft Biometrics fue el último de los experimentos realizados. El objetivo de este experimento era ver la influencia que proporcionaban los Soft Biometrics en el reconocimiento de un individuo.

Para este último experimento y dados los resultados que se habían obtenido hasta ahora con la utilización de las bases de datos, se escogió utilizar la base de datos LFW funneled, dado que era la base de datos que mejores resultados había proporcionado y de la que teníamos más datos almacenados en cuanto a variación de fotos de un mismo individuo.

En la primera parte del experimento se siguió el protocolo experimental utilizando los datos de evaluación de LFW, aplicando el proceso de comparación de caras y obtención de puntuaciones, para este proceso se seleccionaron veinte listas diferentes, cada una con parejas de 300 nombres de imágenes, de esas veinte listas, diez incluían parejas genuinas, es decir, cada pareja pertenecía a la misma persona y las otras diez listas a su vez incluían parejas de impostores, cada foto de la pareja pertenecía a una persona diferente. Se obtuvieron de cada comparación una puntuación global que correspondía a la probabilidad de que las dos fotos pertenecieran a la misma persona. Cada una de las diferentes puntuaciones se almacenó para las siguientes partes del experimento.

Por otro lado, se crearon tres tipos diferentes de vectores de etiquetado de atributos, para cada una de las imágenes de LFW se almacenó en un vector todos sus atributos ordenados. El primero de los vectores estaba formado por los atributos anotados manualmente por una persona al observar cada imagen, etiquetado explicado en la sección 2.4, este etiquetado es el que más atributos tiene almacenados, seis, siendo estos los de género, edad, etnia, sonrisa, barba y bigote. El segundo de los etiquetados estaba formado por los resultados obtenidos del sistema de detección de atributos de Face++, este etiquetado estaba formado por los Soft Biometrics de género, edad, etnia y sonrisa. Por otro lado, el tercero de los etiquetados estaba formado por los Soft Biometrics obtenidos del detector de atributos de Microsoft, y los Soft Biometrics que contenía eran los mismos que el primer etiquetado a excepción de la etnia y con los Soft Biometrics de Beard y Moustache. Tanto en el vector con los datos de Face++

como con los datos de Microsoft todas las imágenes que en su detección de atributos habían tenido múltiple cara o bien no se había detectado cara eran marcadas con un etiquetado especial de error.

Una vez creados los vectores, se realizó la comparación de Soft Biometrics. Para la comparación de Soft Biometrics se utilizaron las mismas veinte listas utilizadas para la comparación de la cara. Como se explicó en el apartado 5.1, el método utilizado para estas comparaciones fue crear un sistema que iba calculando distancia en función de los Soft Biometrics a comparar, para los Soft Biometric de género, barba, bigote, sonrisa y etnia se utilizó la distancia Hamming mientras que para el Soft Biometric de la edad se utilizó la distancia Euclídea. En caso de que alguno de los vectores contuviera una Soft Biometric no detectado esa comparación se desechaba y se pasaba a al siguiente. Posteriormente cada una de las distancias se normalizo de manera inversa entre 0-100, equivalente a la probabilidad de que dos imágenes compartieran el Soft Biometric.

El siguiente paso consistía en la fusión de ambos sistemas, para ello se dividieron las puntuaciones en tres, las obtenidas de la comparación facial dando lugar al vector face, las obtenidas de la comparación de Soft Biometrics, dando lugar al vector Softbio, y la puntuación de fusión, dando lugar al vector Fusión. Para cada una de las comparaciones se realizó una normalización siguiendo estos pasos:

1. Se concatenaron por un lado los vectores dos vectores face existentes, el genuino y el impostor y por otro lado los dos vectores Softbio (genuino e impostor).
2. De cada uno de los vectores concatenados se calculó la media y la desviación típica.
3. Se aplicó la siguiente formula a cada vector:

$$Vector\ normalizado_{SB} = \frac{(Vector\ \text{sin}\ normalizar_{SB} - Media\ vector\ concatenado_{SB})}{Desviación\ típica\ vector\ concatenado_{SB}}$$

$$Vector\ normalizado_{SB} = \frac{(Vector\ \text{sin}\ normalizar_{SB} - Media\ vector\ concatenado_{SB})}{Desviación\ típica\ vector\ concatenado_{SB}}$$

4. Se calculó el vector Fusión gracias a los etiquetados previamente explicados de detección de errores. Para ello se siguieron las siguientes hipótesis:
    a. Si tanto la puntuación face como la puntuación SoftBio procedían de comparación sin errores, el vector fusión estaba formado por la suma aritmética de estas dos puntuaciones.
    b. Si la comprobación Softbio tenía etiquetado previamente un error solo se consideraba la puntuación Face en la fusión, ponderada por un factor 2
    c. Si la comprobación Face tenía etiquetado previamente un error solo se consideraba la puntuación Softbio en la fusión, ponderada por un factor 2
    d. Si en ambas puntuaciones habían etiquetados de error no se consideraban esas puntuaciones para la fusión.

Finalmente se calculó el EER de cada uno de los vectores. Del conjunto de todos los EER obtenidos, diez por cada combinación debido a las diez listas, se extrajo la media y la

desviación típica. Del total de todas las combinaciones posibles se seleccionaron las diez mejores combinaciones que se muestran en las siguientes tabla:

Tabla 6-15: Medias de EER obtenidos del etiquetado de Manual

	Face	Softbio	Fusión
'Gender Age Ethnicity Moustache '	12,70 ± 1,43	14,37 ± 2,65	7,67 ± 1,51
'Gender Age Ethnicity Glasses '		13,23 ± 2,36	8,03 ± 1,31
'Gender Age Ethnicity Beard '		16,03 ± 3,09	8,07 ± 1,52
'Gender Age Ethnicity '		19,10 ± 3,26	8,27 ± 1,24
'Age Ethnicity Glasses Moustache '		18,33 ± 2,88	8,37 ± 1,29
'Age Ethnicity Moustache '		24,73 ± 3,32	8,60 ± 1,38
'Age Ethnicity Beard Moustache '		24,20 ± 3,23	8,60 ± 1,85
'Gender Age Glasses Moustache '		16,30 ± 2,67	8,70 ± 1,32
'Age Ethnicity Glasses Beard '		19,50 ± 2,94	8,73 ± 1,31
'Age Ethnicity Beard '		26,87 ± 3,30	9,03 ± 1,55

Como se puede observar, el sistema de fusión con el etiquetado manual obtiene mejores EERs que el sistema sin la fusión. La mejor combinación sería la formada por los Soft Biometrics de Gender, Age, Ethnicity, Moustache que mejoraría el EER en siete puntos. Pero dado que el etiquetado manual no será el que se implemente en un Sistema real de fusión, veamos los resultados con los otros etiquetados.

Tabla 6-16: Medias de EER obtenidos del etiquetado de Face++

	Face	Softbio	Fusión
'Gender '	12,70 ± 1,43	62,46 ± 2,83	14,05 ± 1,88
'Glasses '		68,37 ± 2,62	15,20 ± 1,28
'Gender Glasses '		41,72 ± 3,58	16,16 ± 2,05
'Age Gender Glasses Ethnicity'		47,46 ± 3,45	16,28 ± 1,48
'Age Gender Glasses '		45,26 ± 2,95	17,46 ± 1,65
'Age Gender Ethnicity'		48,44 ± 2,43	17,49 ± 1,78
'Gender Glasses Ethnicity'		29,25 ± 3,41	17,71 ± 1,53
'Ethnicity'		67,81 ± 2,84	18,82 ± 2,14
'Age Glasses Ethnicity'		49,78 ± 2,64	19,01 ± 1,02
'Gender Ethnicity''		41,96 ± 2,97	19,16 ± 2.02

Tabla 6-17: Medias de EER obtenidos del etiquetado de Face++

	Face	Softbio	Fusión
'Age Gender Glasses Sideburns '	12,70 ± 1,43	39,66 ± 2,14	12,60 ± 1,69
'Age Gender Glasses Moustache '		39,81 ± 2,46	12,87 ± 1,93
'Age Gender Glasses Beard '		39,96 ± 2,82	13,17 ± 1,85
'Age Gender Sideburns '		42,71 ± 4,59	13,39 ± 2,18
'Gender '		62,57 ± 3,32	13,61 ± 1,84
'Age Gender Moustache '		39,27 ± 3,17	13,74 ± 2,19
'Age Gender Glasses '		26,82 ± 3,14	13,89 ± 1,73
'Age Gender Sideburns Moustache '		35,60 ± 3,08	14,19 ± 1,73
'Age Gender Beard '		40,27 ± 3,42	14,49 ± 2,03

Por otro lado, con los etiquetados obtenidos gracias a los estimadores de Soft Biometrics, vemos como con el etiquetado de Face++ con ninguna combinación conseguimos mejorar el EER del sistema son fusión. Mientras que con el etiquetado de Microsoft lo conseguimos mejorar solo con la combinación de Age, Gender, Glasses, Sideburns.

7 Conclusiones y trabajo futuro

El objetivo principal de este Trabajo de Fin de Grado consistía en llevar a cabo un estudio sobre algoritmos automáticos para detección de atributos faciales y Soft Biometrics. Para ello se han seleccionado y estudiado detalladamente dos potentes detectores de atributos como son Face++ y Microsoft Cognitive.

La conclusión más inmediata es que, aunque la estimación de atributos faciales está muy avanzada todavía le queda un largo camino por recorrer. Esto es, aunque ambos sistemas de estimación han demostrado tener altas tasas de éxito en la detección de caras todavía se siguen cometiendo errores graves. Imágenes con caras nítidas se quedan si detectar porque alguna de las herramientas falla en sus algoritmos.

Por otro lado, dentro del análisis individual de Soft Biometrics se ha podido observar cómo el atributo de la edad es con diferencia el más difícil de detectar. Así, mientras que Microsoft detecta correctamente las edades en un 60 % de los casos, Face++ no llega ni al 40 %, siendo ambas cifras muy bajas para un buen estimador. Atributos como el género, la etnia o las gafas parecen ser los que menos problemas generan a la hora de ser estimados. Si bien Microsoft detecta con mayor precisión el género y la edad, Face++ obtiene un mejor rendimiento en gafas y sonrisa.

De la primera parte correspondiente a la detección de atributos puede concluirse que la herramienta de detección de Microsoft ha aportado mejores resultados, dado que demuestra unas mayores tasas de acierto en la estimación en general y dado que su velocidad de detección es casi doscientas veces más rápida que la de Face++.

Finalmente se ha realizado la fusión de los sistemas de reconocimiento facial junto con el reconocedor de Soft Biometrics con el fin de implementar una herramienta de reconocimiento mejorada. Los resultados sugieren que con un buen etiquetado de los Soft Biometrics sería posible desarrollar un sistema de reconocimiento notablemente mejorado, si bien de momento este etiquetado solo se logra de manera manual. Los sistemas de estimación de Microsoft y Face++ todavía tendrían que perfeccionar la estimación para poder implementar una fusión superior.

Por su parte, las líneas de trabajo que se abren con los desarrollos realizados en este Trabajo de Fin de Grado pueden resumirse en dos mejoras para el sistema de fusión:

- Utilizar valores continuos en vez de valores discretos con algún tipo de valor de confianza sobre cuánto de segura es la estimación. Por ejemplo, si en la edad, en vez de utilizar un etiquetado del 0 al 4 utilizamos el propio valor de la edad, obtendríamos unos resultados más reales, aunque también cabe la posibilidad de exponer al sistema a un mayor error en la estimación.
- Por otro lado, tal y como se presentó en la sección 2.3, no todos los Soft Biometrics tienen la misma repercusión a la hora de discriminar y su información no resulta igual de importante. Por consiguiente, la implementación de un sistema de ponderación podría resultar muy útil para la fusión de sistemas.

Referencias

[1] Sergio D.Werner, “Aplicación de Nuevas Tecnologías al Sistema Electoral – Biometría y Voto Electrónico”, 2011

[2] Leer más: http://www.monografias.com/trabajos82/biometria-y-voto-electronico/biometria-y-voto-electronico2.shtml#ixzz4l6GGcUIH

[3] SYRIS Technology Corp, “Technical Document About FAR, FRR and EER”, 2014

[4] Antitza Dantcheva, Petros Elia, Arun Ross, “What Else Does Your Biometric Data Reveal? A Survey on Soft Biometrics”, IEEE Transactions on Information Forensics and Security, 2015

[5] Pedro Tome, Ruben Vera-Rodriguez, Julian Fierrez and Javier Ortega-Garcia, “Facial Soft Biometric Features for Forensic Face Recognition”, Julio 2015

[6] Anil K. Jain, Sarat C. Dass, and Karthik Nandakumar, “Soft Biometric Traits for Personal Recognition Systems”, 2004.

[7] B. Cid Fernández, “Interfaz Gráfica de Etiquetado de Atributos Faciales”, 2016

[8] HD Wallpapers, “http://awallpapersimages.com/2016/07/top-50-emma-stone-hd-images-photos-and-wallpapers-free-download/”

[9] Microsoft, “https://azure.microsoft.com/es-es/services/cognitive-services/face/”

[10] Haoqiang Fan, Zhimin Cao, Yuning Jiang, Qi Yin, Chinchilla Doudou, “Learning Deep Face Representation”

[11] Erjin Zhou, Zhimin Cao, Qi Yin, “Naive-Deep Face Recognition: Touching the Limit of LFW Benchmark or Not?”, 2015

[12] Vitomir Struc, “The INface toolbox v2.0 The Matlab Toolbox for Illumination Invariant Face Recognition”

[13] S. Yang, P. Luo, C. C. Loy, and X. Tang, "From Facial Parts Responses to Face Detection: A Deep Learning Approach", in IEEE International Conference on Computer Vision (ICCV), 2015

Glosario

API	Application Programming Interface
LFW	Labeled Faces in the Wild
EER	Equal Error Rate
CNN	Convolutional Neural Network

Anexo 1: etiquetado obtenido con la interfaz gráfica

En este anexo se puede observar el criterio seguido en la estimación de atributos y Soft Biometrics con la interfaz gráfica de etiquetado manual

- Gender
  - Male→0
  - Female→1

- Age
  - Baby→0 si su edad está entre los 0 y 4 años.
  - Child→1 cuando aparente 17 años o menos.
  - Youth→2 entre los 18 y 39 años.
  - Middle_Aged→3 entre 40 y 60 años.
  - Senior→4 cuando aparente más de 60 años.

- Ethnicity
  - White→0 cuando sea de piel blanca.
  - Black→1 cuando sea de piel negra
  - Asian→2 cuando tengan los ojos achinados
  - Indian→3 cuando tenga rasgos indios.
  - Other_Mixture→4 cuando no tenga claro que pertenezca a las etnias anteriores

- Forehead
  - Fully_Visible→0 Cuando se vea la frente completamente
  - Partially_Visible→1 cuando se vea más o menos la mitad de la frente, ya sea porque tiene gorra o por flequillo
  - Obstructed→ aquella frente que por cualquier razón (pelo u oclusión) está casi o completamente obstruida

- Mouth
  - Open_Widely→0 siempre que los dientes estén separados (los dientes de arriba y los dientes de abajo)
  - Partially_Open→1 boca abierta pero los dientes no separados
  - Close→2 los labios están pegados

- Eyes
  - Open→0 si tiene los ojos completamente abiertos
  - Partially_open→1 si tiene los ojos entre abiertos
  - Close→2 si tiene los ojos cerrados

- Glasses
  - No_Glasses→0 si no lleva gafas
  - Eye_Wear→1 si lleva gafas de ver (con el cristal transparente)
  - Sunglasses→2 si lleva gafas de sol (con el cristal oscuro)

- Smiling
  - Yes→0 si tiene cara de alegría y se le ven los dientes o si tiene cara de alegría, pero no se le ven los dientes
  - No→1 si no se le ven los dientes y está serio

- Oclussion
  - Nose→0 si tiene la nariz más tapada que visible
  - Mouth→1 si tiene la boca más tapada que visible
  - Chin→2 si tiene la barbilla más tapada que visible
  - Left_Ear→3 si tiene la oreja izquierda más tapada que visible
  - Right_Ear→4 si tiene la oreja derecha más tapada que visible

- Beard
  - yes→0 si se ve claramente que tiene barba
  - No→1 si no se ve claramente que tiene barba

- Moustache
  - yes→0 si se ve claramente que tiene bigote
  - No→1 si no se ve claramente que tiene bigote

- Pose
  - Frontal→0 siempre y cuando los dos ojos sean visibles se considerará que la pose del individuo es frontal.
  - Left_Side→1 cuando solo sea visible un ojo y siempre que el lado de la cara que se ve sea el derecho (según el punto de vista del observador)
  - Right_Side→2 cuando solo sea visible un ojo y siempre que el lado de la cara que se ve sea el izquierdo (según el punto de vista del observador)